TOWARD A STOCHASTIC CALCULUS, II*

By E. J. McShane

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF VIRGINIA, CHARLOTTESVILLE

Communicated April 18, 1969

Abstract.—In a preceding note (these PROCEEDINGS, **63,** 275 (1969)) singly and doubly stochastic integrals were defined. Here correspondingly generalized stochastic differential equations are studied. For constructing stochastic models of physical processes with random noises, by proper selection of the doubly stochastic terms, we remove the apparent discordances between classical and stochastic models.

---

1. *Substitution; Differentiation of Composite Functions.*—In this note we shall often write equations involving differentials. We do not actually define differentials; the equations with differentials are merely abbreviations for their integrated forms. Thus a process x is said to satisfy

$$dx^\alpha = f^\alpha\, dt + \textstyle\sum_\rho g_\rho{}^\alpha\, dz^\rho + \sum_{\rho,\sigma} h_{\rho,\sigma}{}^\alpha\, dz^\rho\, dz^\sigma \tag{1.1}$$

if it satisfies, for $a \leqq t \leqq b$,

$$x^\alpha(t) = x^\alpha(a) + \sum_{\rho=1}^{r} \int_a^t g_\rho{}^\alpha\, dz^\rho\,(s) + \sum_{\rho,\sigma=1}^{r} \int_a^t h_{\rho,\sigma}{}^\alpha\, dz^\rho\,(s) dz^\sigma\,(s). \tag{1.2}$$

In classical calculus, expressions involving differentials are usually linearized forms of expressions involving increments. For the stochastic analogues we must also include terms of second degree. We thus have the memory device: Treat differentials dz^ρ as increments, then discard all terms with more than *two* dz-factors and also all those terms with two dz-factors whose doubly stochastic integrals are known to vanish. Thus, if all z^ρ are Lipschitzian, we come back to the familiar linearization of ordinary calculus. Of course the formulas thus obtained must be shown to be valid under suitable continuity conditions.

For example, if $F, f^\alpha, g_\rho{}^\alpha, h_{\rho,\sigma}{}^\alpha$ are processes with property (II) (cf. ref. 4 (3.2)) and x^α ($\alpha = 1, 2$) satisfies (1.1), the memory device gives formulas for $F dx^1$ and $F dx^1 dx^2$, the latter being

$$F(t)dx^1(t)dx^2(t) = \textstyle\sum_{\rho,\sigma} F(t) g_\rho{}^1(t) g_\sigma{}^2(t) dz^\rho(t) dz^\sigma(t), \tag{1.3}$$

that are in fact assertions about stochastic integrals. We can prove these formulas, but only under the extra hypothesis that for $a \leqq s < t \leqq b$ and ρ in $1,\ldots,r$

$$\lim_{t-s=0} \text{ess. sup. } E([z^\rho\,(t) - z^\rho\,(s)]^4 | \mathfrak{F}_s)/(t-s) = 0 \tag{1.4}$$

uniformly on $[a, b]$.

For another example we have a generalization of the Itô differentiation formula:[5]

THEOREM 1.1. *Let $F = (F(x,t)\colon\ x \epsilon R^n,\ t \epsilon [a,b])$ be continuous together with F_t, F_{x^i} and $F_{x^i x^j}$ $(i,j = 1,\ldots,n)$. Let $z^1,\ldots,z^\rho$ satisfy (2.1) of reference 4 and also, for $a \leqq s < t \leqq b$ and $\rho \epsilon \{1,\ldots,r\}$,*

$$\lim_{t-s\to 0} \text{ess. sup. } E([z^\rho(t) - z^\rho(s)]^c \mathfrak{F}_s)/(t-s)^2 = 0 \tag{1.5}$$

uniformly on $[a,b]$ *for* $c = 6,8$. *Let* $x^i, f^i, g_\rho{}^i, h_\rho{}^i{}_{,\sigma} (i = 1,\ldots,n; \rho,\sigma = 1,\ldots,r)$ *be processes satisfying (1.1) and having property (II) (cf. ref. 4, (3.2)). Then for t in* $[a,b]$

$$\begin{aligned} dF(x(t),t) = {} & [F_t + \sum_{i=1}^{n} F_{x^i} f^i(t)]dt \\ & + \sum_{i=1}^{n}\sum_{\rho=1}^{r} F_{x^i} g_\rho{}^i(t) dz^\rho \\ & + \{\sum_{i=1}^{n}\sum_{\rho,\sigma=1}^{r} F_{x^i} h^i{}_{\rho,\sigma}(t) \\ & + 1/2\sum_{i,j=1}^{n}\sum_{\rho,\sigma=1}^{r} F_{x^i x^j} g_\rho{}^i(t) g_\sigma{}^j(t)\}\, dz^\rho\, dz^\sigma, \end{aligned} \tag{1.6}$$

where F_t, etc., are evaluated at $(x(t),t)$.

The proofs of these statements are simple in plan but tedious in execution. Given Π, in Theorem 1.1 we expand the differences of $F(x(t),t)$ at successive t_j by Taylor's theorem and approximate the Δx^i by Riemann sums deduced from (1.2). The result consists of the terms needed for (1.6), together with a multitude of unwanted terms with too many factors $\Delta_j z^\rho$. By judicious grouping and occasional reference to Lemma 3.1 of reference 4, it can be shown that the sum of the unwanted terms tends to 0 with mesh Π.

2. *Differential and Functional Equations.*—In reference 2 an existence theorem was proved for solutions of a class of stochastic functional equations. With only trivial changes, this theorem extends to equations containing both singly and doubly stochastic terms. For brevity, we state the result only for stochastic differential equations:

THEOREM 2.1. *Let (2.1) of reference 4 be satisfied, and let* x_0 *be* $\mathfrak{F}_a$*-measurable and have finite second moments. Let* $f^i, g_\rho{}^i, h_{\rho,\sigma}{}^i$ $(i = 1,\ldots,n; \rho,\sigma = 1,\ldots,r)$ *be functions defined and Lipschitzian on* R^n $[a,b]$. *Then the equations*

$$\begin{aligned} dx^i = {} & f^i(x(t),t)dt + \sum_{\rho=1}^{r} g_\rho{}^i(x(t),t)dz^\rho \\ & + \sum_{\rho,\sigma=1}^{r} h_{\rho,\sigma}{}^i(x(t),t)dz^\rho\, dz^\sigma \quad (i = 1,\ldots,n) \end{aligned} \tag{2.1}$$

with initial condition $x^i(a) = x_0{}^i$ *have a solution; and if* x_1, x_2 *are solutions, then for each t we have* $x_1(t) = x_2(t)$ *a.s.*

From reference 2 we can also adapt the method of approximating solutions of (2.1) by "Cauchy polygons." Let Π be a partition of $[a,b]$, with division points $t_1,\ldots,t_{m+1}$. We define $x_\Pi(a) = x(a)$, and successively

$$\begin{aligned} x_\Pi{}^i(t_{j+1}) = {} & x_\Pi{}^i(t_j) + f^i(x_\Pi(t_j),t_j)\Delta_j t + \\ & \sum_{\rho=1}^{r} g_\rho{}^i(x_\Pi(t_j),t_j)\Delta_j z^\rho + \sum_{\rho,\sigma=1}^{r} h_{\rho,\sigma}{}^i(x_\Pi(t_j),t_j)\Delta_j z^\rho \Delta_j z^\sigma; \end{aligned} \tag{2.2}$$

inside each $[t_j,t_{j+1}]$ we define $x_\Pi{}^i(t)$ by linear interpolation.

3. *Stochastic Models.*—Let us assume that a system subjected to Lipschitzian disturbances is known by the theory governing such systems to satisfy

$$dx^i = f^i(x(t),t)dt + \sum_{\rho=1}^{r} g_\rho{}^i(x(t),t)dz^\rho. \tag{3.1}$$

Then for Lipschitzian disturbances it is equally true that it satisfies (2.1) with any $h_{\rho,\sigma}{}^i$, since the doubly stochastic integral of the last term is then 0. But this equivalence of all systems (2.1) fails when we allow disturbances satisfying (2.1) of reference 4, and for a stochastic model permitting such disturbances we must select one system. We can avoid the confusion associated with the choice $h_{\rho,\sigma}{}^i = 0$ by using the following rule.

(3.2) *Selection principle:* When a scientific theory asserts that a system affected by Lipschitzian disturbances $z^1,\ldots,z^r$ satisfies (3.1), then in order to form a stochastic model adequate for use with any disturbances satisfying (2.1) of reference 4 we replace (3.1) by the equation (2.1) (equivalent to (3.1) for Lipschitzian disturbances) in which we define

$$h_{\rho,\sigma}{}^i\,(x,t) = (1/2)\sum_{h=1}^{n} g_\rho{}^i{}_{,x^h}\,(x,t)g_\sigma{}^h\,(t)\ (i = 1,\ldots,n). \tag{3.3}$$

It should be observed that we are not adding a "correction term" to the equation (3.1) valid for Lipschitzian disturbances. Rather, we are regarding (2.1) with notation (3.3) as being the equation of the system for all disturbances z^ρ satisfying (2.1) of reference 4, and are merely refraining from making in all cases the simplification to (3.1) which is valid in the special case of Lipschitzian disturbances, but is not valid in other cases.

The justification of the selection principle must consist of an exhibition of systems in which troubles are present when (3.1) is used and absent when (2.1) (with (3.3)) is used. For reasons of space we present only two examples, one banal.

If a system satisfies

$$dx^1 = 2x^2dz,\ dx^2 = 1dz \quad (x^1(a) = x^2(a) = 0) \tag{3.4}$$

with Lipschitzian disturbance $z(t)$, it is easy to see that

$$x^1(b) = z(b)^2 - z(a)^2 \quad (b > a). \tag{3.5}$$

This fails if z is a Wiener process (ref. 1, p. 444). But if we use the selection principle (3.2), we replace (3.4) by

$$dx^1 = 2x^1dz + (dz)^2, dx^2 = dz \quad (x^1(a) = x^2(a) = 0), \tag{3.6}$$

which is equivalent to (3.4) when z is Lipschitzian. The solution of this by Cauchy polygons (2.2) is trivially easy; we obtain (3.5) as the solution no matter what the disturbance z may be.

Before presenting the next example we note that (2.1) can be simplified in appearance by introducing functions z^{r+1} and x^{n+1}, both identically t, and corresponding new coefficients, such as $g_{r+1}{}^i = f^i$. The new equation has no dt term. Applying the selection principle to the new equation apparently yields a different result; doubly stochastic terms with dz^{r+1}-factors will be present. But these may be discarded, since their doubly stochastic integrals vanish.

If the z^ρ are independent Wiener processes, equations (3.1) are the Itô-Nisio equation. Wong and Zakai[6] suggest that another approach to stochastic differential equations is to replace the z^ρ by continuous approximations $z_\Pi{}^\rho$ that coincide with the z^ρ at the division points of a partition Π and are linear on the intervals of Π, solve the resulting ordinary equations, and then let mesh Π tend to 0. As they say, experiments correspond to such an approach. They proved for $n = 1$, and conjectured for all n, that the solutions corresponding to the z^ρ would converge to the solutions, not of the Itô-Nisio equation with which we started, but of another equation. In view of Examples 1 and 3 of reference 4 this other equation is (2.1) with (3.3), simplified in appearance because the z^ρ are Wiener processes. Thus use of our selection principle takes us by another path to the goal of Wong and Zakai.

However, an even closer relationship exists. If Π divides $[a,b]$ into equal subintervals, all z^ρ except those in a set of arbitrarily small measure satisfy a Hölder condition

$$|z^\rho(t) - z^\rho(s)| \leqq K(t - s)^{.4},$$

so for these z_Π satisfies a Lipschitz condition of constant $K(\text{mesh } \Pi)^{-.6}$. If we now compare the Cauchy polygon for (2.1) with (3.3), having vertices at the division points of Π, with the solution of (3.1) having $z_\Pi{}^\rho$ in place of z^ρ, we find that the agreement is quite good, because (3.3) yields essentially the same improvement in truncation error as the Runge-Kutta method does. It can in fact be shown that the distance between them, in the metric of convergence in probability, tends to zero. Hence, except that we have used subdivisions of $[a,b]$ into equal parts and replaced L_2-convergence by convergence in measure, we have verified the conjecture of Wong and Zakai.

4. *Remarks.*—(1) By means of Corollary 4.1 of reference 4 we can extend the stochastic integrals (as in ref. 5) to integrands $f(t,\omega)$ separable and $F_\cdot$-measurable and a.s. square-summable over $[a,b]$. In fact, for the doubly stochastic integral this holds with "square" omitted.

(2) As indicated in reference 3, it is highly desirable to use that version of the integral or solution that coincides with the pointwise limit at those ω for which the $z^\rho(\cdot,\omega)$ are Lipschitzian. But in the present setting this is the natural thing to do.

(3) In reference 3 we studied continuity of solutions, to be sure that small causes such as rounding did not produce excessive effects. The same techniques yield the same estimates here, but regrettably no better ones.

* This research was supported in part by the Army Research Office grant ARO-D-978.

1 Doob, J. L., *Stochastic Processes* (New York: John Wiley, 1953).

2 McShane, E. J., "Stochastic integrals and stochastic functional equations," *SIAM J. Appl. Math.*, **17**, 287–306 (1969).

3 McShane, E. J., "Stochastic functional equations: Continuity properties and relation to ordinary equations," in *Proceedings of a Workshop on Calculus of Variations and Optimal Control Theory*, ed. A. V. Balakrishnan, (New York: Academic Press, 1969).

4 McShane, E. J., "Toward a stochastic calculus, I," these PROCEEDINGS, **63**, 275 (1969).

5 Skorohod, A. V., *Studies in the Theory of Random Processes* (translated from the Russian by Scripta Technica, Inc., Reading, Mass.: Addison-Wesley, 1965).

6 Wong, E., and M. Zakai, "On the relation between ordinary and stochastic differential equations," *Intern. J. Eng. Sci.*, **3**, 213–229 (1965).